

パラ言語情報の生成と知覚

Production and Perception of Paralinguistic Information

独立行政法人国立国語研究所研究開発部門

Dept. Language Research, National Institute for Japanese Language

前川喜久雄

Kikuo Maekawa

< 研究協力者 >

早稲田大学 Waseda University

菊池英明 Hideaki Kikuchi

国立国語研究所研究開発部門, NIJL

籠宮隆之 Takayuki Kagomiya

国立国語研究所研究開発部門, NIJL

藤本雅子 Masako Fujimoto

金沢工業大学 Kanazawa Institute of Technology

北川智利 Norimichi Kitagawa

The goal of the current study consisted in clarifying the mechanisms of the production and perception of ‘paralinguistic’ information. By paralinguistic information (PI, hereafter) is meant information like speaker’s attitudes and intentions. Although, these informations are essential in phonetic communication, they are rarely studied from a scientific point of view. Production mechanism of PI was studied at the levels of acoustics and physiology. In acoustic studies, six PI types (neutral, focused, disappointment, admiration, indifference, and suspicion/disbelief) were investigated. It turned out that the choice of PI types influenced systematically and strongly all acoustic features investigated, viz., segmental and overall duration, speech fundamental frequency contour, formant frequency of vowels, and spectral tilt of voice source. These findings were applied for the proposal of new intonation labeling scheme that we called X-JToBI. Physiological investigation using EMA revealed that the shift in formant frequencies was caused by the shift in the horizontal positioning of the tongue body and the lip distance. High-speed video imaging of larynx revealed systematic change of phonation types due to PI types, which was congruent with the observed differences of spectral tilt. In perception studies, part of the speech samples used in the acoustic investigation was used as speech samples. First, simple identification task revealed that PI was perceived by native (Japanese) speakers with higher-than 80% accuracy. Second, MDS analysis of the confusion matrix obtained in the identification task revealed that the perceptual space of PI consisted of 3 dimensions. Regression analyses revealed high correlations between the acoustic parameters and MDS dimensions. Third, the same identification task followed by MDS analysis was conducted using non-native (English speaking) subjects in order to examine the language-(in)dependence of PI. The result suggested that the perception of some PI, especially those who were phonetically realized by the choice of phrase-final boundary pitch movements, was very difficult for non-native speakers who knew nothing about Japanese. Lastly, a conceptual model of the production process of PI was proposed based upon the whole results. This model put emphasis upon the distinction between two sorts of PI: the one expressed in terms of the manipulation of phonological features like boundary pitch movement, and the other expressed in terms of the manipulation of voice-quality features.

Key Words: Paralanguage, Paralinguistic information, X-JToBI, EMA, phonation, MDS, language-dependency, voice-quality

1. 研究の目的

藤崎の分類によれば音声に含まれる情報は、言語情報、パラ言語情報、非言語情報の三種に分類される[1]。このうちパラ言語情報は、話者が意図的に表出する情報でありながら文字には転写されない情報であり、話し言葉の本質に深く関係している[2]。本研究ではパラ言語情報の伝達メカニズムを生成と知覚の両面から実験的に検討した[3-6]。

2. パラ言語情報の生成

2.1 音響的分析

2.1.1 音響分析用資料

キャリア文「～ですか」の文頭に「そう」「やまの(山野)さん」「あなた」「え(絵)」「えー(A)」(記号はアクセント核)等15語を配した発話を、標準語話者3名に種々のパラ言語的意味を意図して繰り返し(最低10回以上)発話させた資料を分析する。パラ言語的意味としては文献[7]を参考として「N中立」「F強調」「I無関心」「A感心」「D落胆」「S疑い」の6種類を採用した。話者3名は男性2名(ST,YS)、女性1名(JH)である。

2.1.2 同定実験

最初に資料の妥当性を検討する。全資料のうち「そうですね」「山野さんですか」「あなたですか」を選び全話者の全436発話に対し上記6種類のパラ言語情報の同定実験を実施した。被験者は20名である。表1に実験結果を示す。「F」以外の情報は0.8以上の正確率で知覚されており、就中「A, D, S」の正解率が高い。

	A	D	F	I	N	S
A	0.89	0.01	0.09	0.01	0.00	0.00
D	0.01	0.99	0.00	0.00	0.00	0.00
F	0.01	0.00	0.59	0.05	0.34	0.00
I	0.01	0.03	0.01	0.81	0.14	0.00
N	0.00	0.02	0.05	0.07	0.86	0.01
S	0.00	0.01	0.01	0.00	0.00	0.98

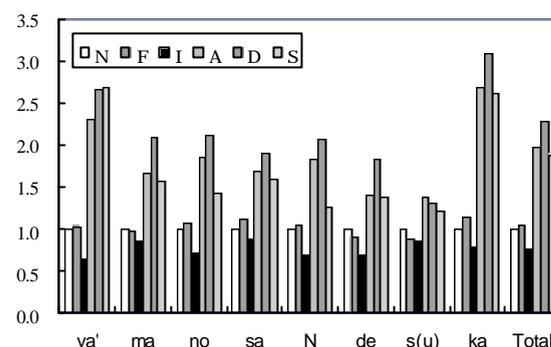
表1. パラ言語情報の同定実験結果

以下の音響分析では同定率が0.5以上であったサンプルのみを対象とした。

2.1.3 持続時間の特徴

まず発話全体の持続時間を検討する。図1は「山のさんですか」全体の持続時間(右端)と発話を構成する各モーラの持続時間平均値がパラ言語的の情報による変動する様子を「N」の平均を1.0として表示している。

発話全体としても各モーラにおいても「N, F, I」は短く「A, D, S」は長いことがわかるが、同時に発話冒頭と



発話末尾のモーラにおいては、変動が殊に著しい。

図1. パラ言語情報による発話持続時間の変動(話者ST)

図1では「A, D, S」の平均持続時間長が「N」の2倍を越えている。このような変動下において、日本語の重要な音韻特徴である分節音の長短の対立が保存されるかどうかを検討した結果を図2に示す。

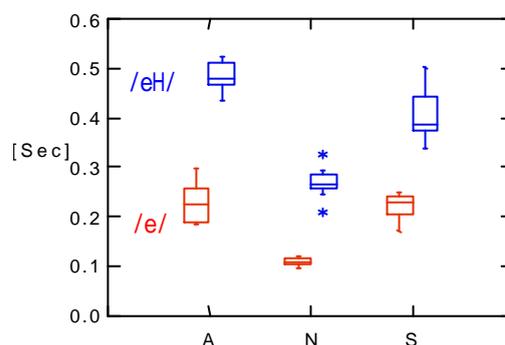


図2. 短母音(赤)と長母音(青)の持続時間長(話者ST)

図2は「絵ですか」と「Aですか」の発話冒頭に位置する短母音/e/と長母音/eH/の持続時間平均値の差が「A, N, S」の変動下において明瞭に保存されていることを示している。

2.1.4 音声基本周波数(F0)の特徴

図3に発話「そうですね」におけるF0曲線の典型例を示す。第3音節「す」では母音が無声化するためにF0が観察されないことに注意。図中の縦線はスペクトログラムから決定した音声境界、三角は「そう」のアクセント核に起因するF0の下降開始時刻を示している。「N, F, I」の形状は、この発話に対して音韻論が予測するF0形状に合致しているが、「A, D, S」ではいくつかの点において逸脱が生じている。

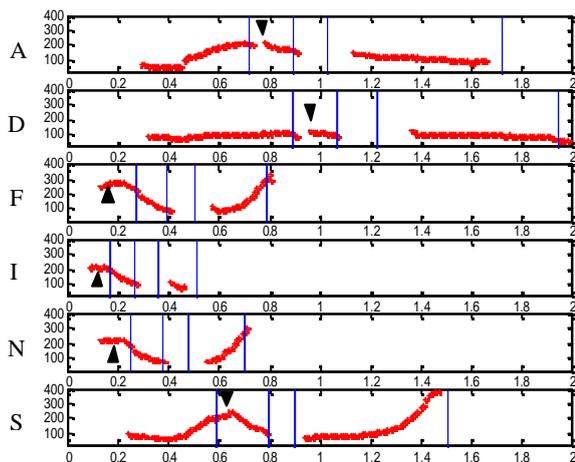


図3.「そうですか」の典型的F0形状(話者ST)

2.1.5 句頭の上昇

図3の発話は重音節で始まるため、通常は句頭におけるF0上昇が弱化するか消失する。実際図3のうち「N, F, I」はそのようであるのだが、「A」と「S」では明瞭な上昇とそれに先立つ低ピッチ区間の延長が生じている。図4に上昇幅の平均値の分布を示す。

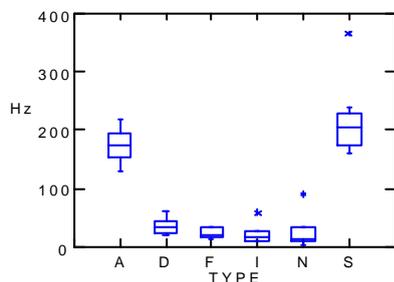


図4.「そうですか」における句頭F0上昇幅の比較(話者ST)

2.1.6 句末の上昇

図3では「N, F, S」の句末に上昇イントネーションが生じているがその上昇形態は「N, F」と「S」とで質的に異なっている。前者が単純な上昇調であるのに対し、後者ではまず低ピッチ区間が持続した後に上昇が始まる。また上昇の幅も大きい。これは川上が文献[8]で夙に指摘した4種の上昇調のうち「反問の上昇」に該当する。

2.1.7 アクセントによる下降のタイミング

最後に、アクセント核によるF0下降のタイミング(図3の三角記号参照)にもパラ言語情報の影響が認められる。図5に「あなたですか」における下降タ

イミングの平均値の分布を示す。「A, S」の平均値には「N, F, I」に比べると遅れが生じている。

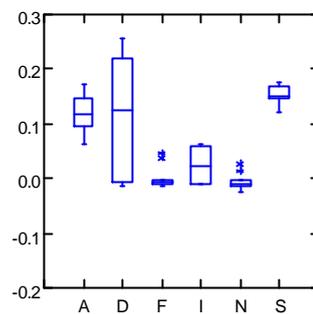


図5.アクセント核によるF0下降タイミングの分布(話者STの「あなたですか」)

2.1.8 フォルマント周波数

パラ言語情報の影響はいわゆる分節的特徴にも及ぶ。図6では「あなたですか」に含まれる4個の/a/について第1第2フォルマント周波数(F1, F2)の散布図を示した。「S」のF2は組織的に「A」よりも上昇している。

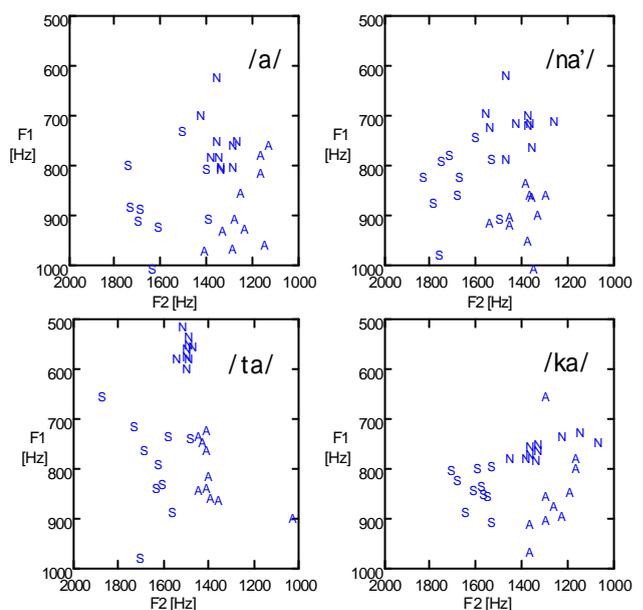


図6.「あなたですか」に含まれる/a/のF1-F2平面での分布(話者ST)

2.1.9 スペクトル傾斜

ここまで利用してきた音声資料を聴取すると発声様式(phonation type)の差に気づく。「D」には明瞭な息漏れ声、「A」はそれよりも弱い息漏れ声、そして「S」では強い発声(pressed phonation)がおこなわれているように聞こえる。この印象を定量的に評価するためにスペクトル傾斜の評価をおこなった[9]。図7は1024点

FFTにおけるH1-A3(第1倍音と第3フォルマント周波数に最も近接した倍音のレベル差)である。分析対象としたのは次節で言及するEMAによる調音運動測定実験のために録音された音声のうち、「笹田が」に含まれる4個の/a/であり、図7はすべての母音をプールした場合の結果である。「D」の傾斜が最も大きく、「S」の傾斜が最も小さいことは聴覚的印象からの予測と一致している。また「N」よりも「A」の方が「D」に近い値をとっていることも予想に合致する。

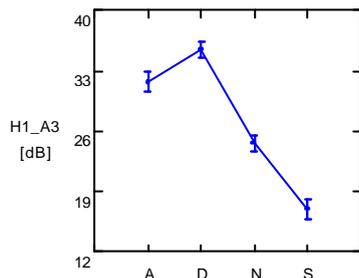


図7. スペクトル傾斜の測定結果(話者 ST)
単位は[dB]。図中の縦線は標準偏差。

2.2 調音運動の分析

2.2.1 舌の位置

2.4節で報告したフォルマント周波数の差が音響分析の誤差ではなく、実際に調音運動の差によって生み出されていることを確認するためにEMA装置による調音運動の測定を実施した。舌、上下唇、下顎に配した6個の測定用コイルのうち、舌背に対応するT3の測定結果を図8に示す。

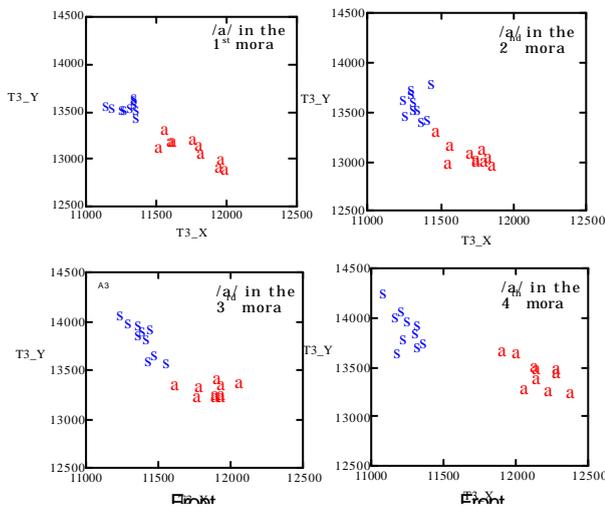


図8. EMAによる舌背位置の測定結果(話者 ST)
横軸が水平位置、縦軸が上下位置。単位は 10^2 mm。

「S」のサンプル群は明らかに「A」よりも左側(すなわち声道の前寄り)に偏差して位置していることがわかる。図9は「笹田が」に含まれるすべての分節音に対してT3コイルの前後(水平方向)位置の平均値と標準偏差を示したものである。図8と同じ差異がすべての子音(/s,d,g/)に観察される。

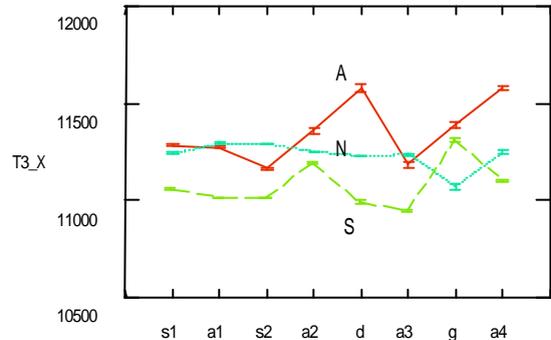


図9. 舌背位置の平均値(話者 KM)。横軸が水平位置、縦軸が上下位置。単位は 10^2 mm。

以上の結果から図6に観察されるF2の差異の背景には舌の調音運動の差異が実際に存在していることを証明している。なお同じEMAデータの解析により両唇の調整もF2の調整に関係していることが示唆されている[3]。

2.2.2 声門面積

図7が示唆する発声様式の差異の直接的証拠を得るために声帯振動の画像解析を実施した。鼻腔から上咽頭を介して喉頭直上までファイバースコープを挿入し毎秒4500フレームの画像を撮影した。この実験では撮影時間の制約(最大0.7秒)により「駅」「秋」などの一語文を用い、「N, D, S」発話時の声門の状態を記録した。図10は発話「駅」中の母音/e/の末尾約200msにおける声門開大面積の変化を示している。

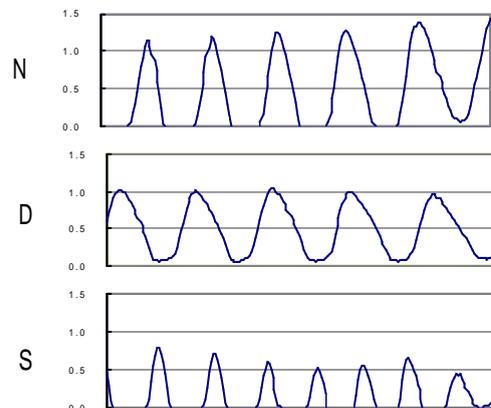


図10. 声門面積の変化パターン(話者 KM)。縦軸の単位は任意。

「D」では声門が完全には閉鎖されておらず、息漏れ発声を証拠だてている。一方「N,S」では声門は完全に閉鎖されている。「S」では声帯振動の周期に占める閉鎖区間の比率が「N」よりも大きく、「S」における pressed phonation を証拠だてている。

2.3 韻律ラベリングへの応用

2.1節に報告した F0関係の特徴はパラ言語情報を含む音声の韻律ラベリングに貴重な示唆を与える。本研究と同時期に進行していた『日本語話し言葉コーパス』の韻律ラベリング手法である X-JToBI には本研究の成果が直接に応用されている。紙幅の関係で解説は省略に従うので、詳細については文献[10]を参照されたい。

3. パラ言語情報の知覚

3.1 パラ言語情報の知覚空間

多次元尺度法(MDS)を用いてパラ言語情報の知覚空間を構築することを試みた。MDSの入力となる類似性行列は表1に示した同定実験データから以下の方法で計算した。任意の2刺激間の類似性をそれらが同一のパラ言語情報として同定された確率と定義する。その際6種類の情報のどれに同定されたかは問わない。表2の場合被験者10名中2名が刺激I,Jを「N」と同定し、3名が「D」、1名が「A」と同定しているので、刺激IとJの類似性は $6 \div 10 = 0.6$ となる。この計算を2.1.2で説明した436サンプルに対して実施し436行×436列の類似性行列を得た。これには音響分析の対象外とした同定率50%以下のサンプルも含まれていることに注意。

表2. 類似性の計算例

STIMULI	SUBJECTS									
	1	2	3	4	5	6	7	8	9	10
I	N	S	D	F	N	D	D	A	S	D
J	N	N	D	D	N	D	A	A	D	D

SASのMDSプロシージャを用いて分析したところ、3次元解で STRESS=0.04となった。図11に刺激の散布状態を示す。この図では個々のサンプルの分布は省略して、各被験者(3名)についてパラ言語情報ごとに68%確率楕円を表示している。図中のシンボルは、話者者STによる「N」の分布を“st-N”と示している。

図11では、話者YSとJHによる「N」と「F」が一部重複している以外は、すべてのパラ言語情報が明瞭に分離している。また異なる話者によって発音されたサンプルもパラ言語情報が同一であれば同じ領域に分布

している。

図11の第1次元(DIM1)は「A,D,S」と「N,F,I」を分離している。「A,D,S」は持続時間が長く、振幅が大きく、F0と発声様式にも特徴が認められるなど、音声学的に顕著な性格を有するサンプルであることから、DIM1はパラ言語情報の「顕著さ(salience)」の軸と解釈できる。第2次元(DIM2)は「S」と「A,I」を分離している。「S疑問」が聞き手の反応(回答)を要求する発話である一方、「A」や「I」が反応を要求しない発話であることから、DIM2は「問合せ(query)」の軸であると解釈できる。第3次元(DIM3)は「大きさ(loudness)」の軸と解釈できる。聞こえの大きなサンプル「A,F,S」と聞こえの小さなサンプル「N,I,D」を分離しているからである。

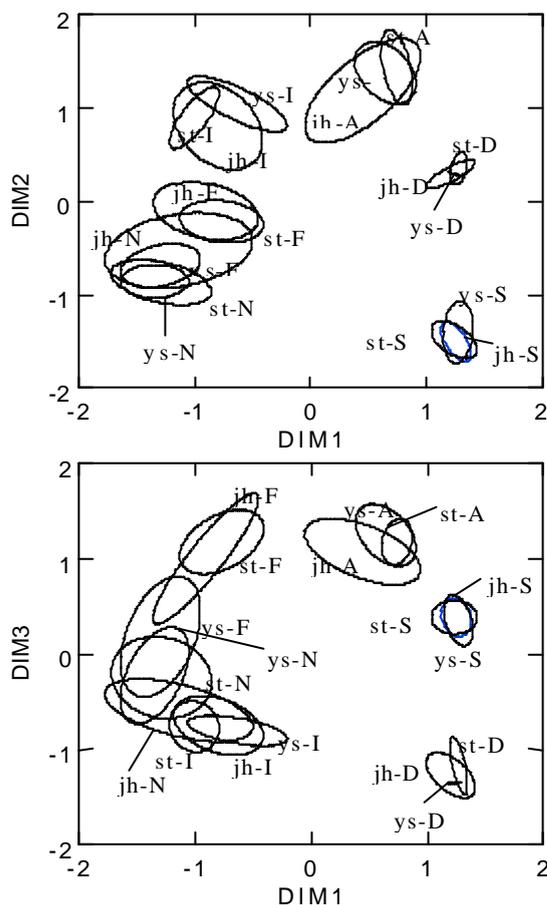


図11. MDSにより構成されたパラ言語情報の知覚空間

3.2 音響特徴との関係

MDS空間における布置とサンプルの音声学的特徴との関係を検討するために、各サンプルのDIM1-3上での座標値を音響分析で得られた種々の特徴量との回帰

分析を実施した。独立変数として用いたのは、発話全体の F0の平均値 (F0)、F0 レンジの平均値 (PR) RMS 振幅平均値 (RMS) および持続時間長平均値 (DUR) である。発話全体とは別に発話冒頭および末尾の 2 モーラについて上記の平均値を計算した。これらを例えば "F0_I", "F0_F" のように添え字で区別する。"PeakTiming" は図 3, 5 で検討したアクセント下降タのイミングであり、F1_Dev, F2_Dev, F3_Dev はそれぞれ第 1、第 2、第 3 フォルマント周波数の平均値からの偏差である。発話者 JH (女性) の発話は F0 が非常に高いために安定したフォルマント推定が困難であったために測定から除外した。

表 3. 重回帰分析の結果。すべての独立変数を利用し話者 2 名の場合。各セルの数値は標準化偏相関係数。ゴシックおよび下線を施した数値はそれぞれ 1% および 0.1% 水準で有意。最下行は重相関係数。

INDEPENDENT VARIABLES	DEPENDENT VARIABLES		
	DIM1	DIM2	DIM3
F0	0.117	0.037	<u>0.299</u>
F0_I	<u>-0.210</u>	0.049	0.149
F0_F	0.035	-0.139	-0.137
PR	-0.034	<u>0.470</u>	0.208
PR_I	-0.013	-0.010	<u>0.302</u>
PR_F	-0.044	<u>-1.047</u>	-0.046
RMS	-0.098	0.051	<u>0.497</u>
RMS_I	0.045	0.153	-0.211
RMS_F	0.081	0.080	0.069
DUR	<u>0.616</u>	0.047	<u>0.173</u>
DUR_I	0.041	-0.083	0.168
DUR_F	<u>0.408</u>	<u>0.371</u>	<u>-0.208</u>
Peak Timing	<u>0.204</u>	<u>0.180</u>	<u>-0.180</u>
F1_Dev	0.046	0.011	0.061
F2_Dev	<u>0.103</u>	<u>-0.242</u>	<u>-0.286</u>
F3_Dev	-0.009	0.036	0.144
R ²	0.892	0.601	0.642

表 4. 重回帰分析の結果。フォルマント周波数以外の独立変数を利用し話者 3 名の場合。表示法は図 3 と同じ。

INDEPENDENT VARIABLES	DEPENDENT VARIABLES		
	DIM1	DIM2	DIM3
F0	0.217	0.032	<u>0.671</u>
F0_I	<u>-0.254</u>	0.003	-0.152
F0_F	0.032	-0.147	<u>-0.359</u>
PR	0.027	<u>0.415</u>	<u>0.293</u>
PR_I	0.035	0.119	-0.066
PR_F	-0.071	<u>-1.050</u>	0.054
RMS	-0.059	0.009	<u>0.625</u>
RMS_I	0.058	0.106	-0.186
RMS_F	0.057	0.127	0.109
DUR	<u>0.598</u>	0.020	<u>0.198</u>
DUR_I	<u>0.116</u>	-0.075	0.139
DUR_F	<u>0.425</u>	<u>0.374</u>	-0.137
Peak Timing	<u>0.264</u>	<u>0.169</u>	<u>-0.212</u>
R ²	0.876	0.645	0.519

表 3 にすべての独立変数を用いた重回帰分析の結果 (話者は 2 名) を、表 4 にフォルマント周波数を除外

した分析の結果 (話者は 3 名) を示した。表 3, 4 とともに DIM1 は DUR と、DIM2 は PR と、そして DIM3 は RMS と相関していることがわかり、各軸の解釈と一致する。また F0, PR, RMS, DUR, Formant (F2) のすべてに有意性が認められたことも注目に値する。

3.3 非母語話者による知覚

前節と同じ実験を日本語を母語としない被験者群に実施することによって、パラ言語情報の言語依存性を検討した [11]。被験者としては、日本語の学習経験があり実験時に日本に滞在していた米語母語話者 (語学教師ないし情報処理技術者) 11 名と、日本語学習経験が皆無の米語話者 (州立大学の学部学生) 15 名の 2 群を用いた。前者を学習者群、後者を非学習者群と呼ぶ。学習者群には日本語能力検定試験 1 級合格者が 2 名含まれているが、それ以外の被験者の日本語は中級以下レベルである。知覚実験の方法は日本語母語話者の場合と全く同一である。MDS 分析における STRESS 値 (3 次元解) は、学習者群で 0.07、非学習者群で 0.16 となった。図 12 に学習者群、図 13 に非学習者群における布置を示す。図 11 におけるような発話者の明瞭な分離が認められなかったため、図 12, 13 の確率楕円では発話者を区別せず、パラ言語情報のタイプだけを表示している。

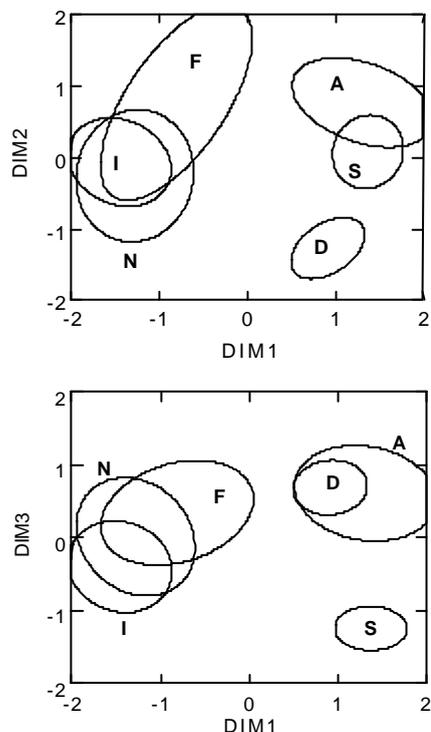


図 12. MDS により構成されたパラ言語情報の知覚空間
米語を母語とする日本語学習者群の結果

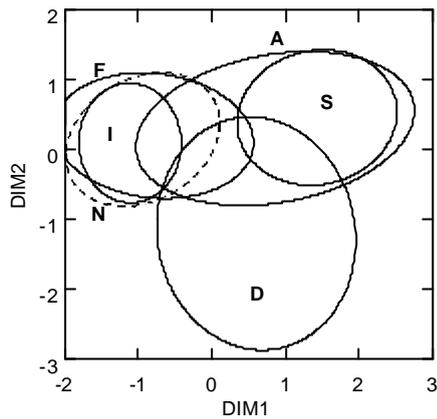


図13. MDSにより構成されたパラ言語情報の知覚空間
米語を母語とする日本語非学習者群の結果

学習者群の知覚空間は母語話者群の空間に類似している。DIM1は「顕著さ」の軸と解釈でき、また DIM2と DIM3はそれぞれ母語話者群の DIM3と DIM2に対応しているとみなせる。ただし刺激の弁別は十分でなく、DIM2では「A」と「S」、「N」と「I」の分布に重複が生じ、また DIM3では「D」の分布が「A」のなかにとりこまれている。

非学習者群の知覚空間は一見すると母語話者群とは非常になった構造をもっている。この群の DIM3ではすべてのパラ言語情報の分布が大幅に重複しており、無意味な軸となっている（そのため図13では DIM3を表示していない）。

しかし非学習者群の DIM1は、母語話者群と同様（従って学習者群とも同じく）「A,D,S」と「N,F,I」を分離しており、やはり「顕著さ」の軸と解釈することができる。また、DIM2は「D」とそれ以外のパラ言語情報とを分離する軸となっているが、この構造は、実は母語話者群の DIM3および学習者群の DIM2が果たしている機能でもある。図11の DIM1-DIM3のパネルないし図12の DIM1-DIM2のパネルを時計回りに45度回転させると、図13に類似した分布が得られる。非学習者は母語話者および学習者とは少し異なった形で「大きさ」の軸を保有していると考えられる。

それでは母語話者ないし学習者における「問合せ」の軸はどうかというと、図13からこれに該当する軸を読みとることができない。実際図13では DIM1においても DIM2においても「S」と「A」とがほぼ完全に重複してしまっている。この点が非学習者群の知覚空間の大きな特徴である。

4 議論

4.1 パラ言語情報の言語依存性

最初に3節の実験結果について議論する。非学習者群の知覚空間には母語話者群の DIM2（学習者群の DIM3）である「問合せ」の軸が欠落していた。母語話者にとってこの軸を認知するための重要な手がかりとなるのは発話末尾のイントネーション特徴である。これは表2,3においては DIM2に対する変数 PR_F の高い寄与となって表現されている。

非学習者群が日本語における二種類の上昇イントネーションを聞き分けることができない（あるいは聞き分けることができても、その機能を正しく推定することができない）ことはパラ言語情報に一定範囲で言語依存性が認められることを示している。この発見にはふたつの意義がある。第一に、言語情報が完全に言語依存性の情報であること、また非言語情報である感情の知覚にはかなり高度な言語普遍性が認められること [13]と比較すると、パラ言語情報が両者の中間的性格をもつことが示唆される。第二に、この発見は今回のサンプルに含まれていた二種類の上昇調イントネーションの間には明瞭な記号的差異が存在していることの強い証拠と解釈することができる。

もちろん、以上の議論を裏返せば、学習者群および非学習者群が（程度の差はあれ）母語話者群と同様に知覚することができた軸（母語話者の DIM1と DIM3）に関しては、パラ言語情報の言語普遍性を認めることになる。

4.2 音声生成におけるパラ言語情報の位置

本研究で得た知見に基づき、パラ言語情報と非言語情報（感情や話者の個人性）を考慮に入れた音声生成過程の概念モデルを提案する（図14）。本モデルの特徴はパラ言語情報に関する発話のプランニングにふたつの異なる経路を認める点にある。すなわち言語情報を参照するプランニングと、参照なしのプランニングである。図2に示した母音長の音韻対立を保存した発話持続時間長の制御や二種類の上昇調イントネーションの選択は前者に属する。またアクセント核による F0 下降タイミングの制御も、それが知覚されるアクセント位置の変更を伴わない範囲で制御されているという点で前者に属する制御であると考えられる。これら以外のすべての制御（発話全体の持続時間、発話全体の

F0レンジ、発話全体の振幅、発声様式など)が後者に属する。図14ではこれらを声質(Voice-quality)と総称している。

一方、非言語情報のプランニングの特徴は言語情報を参照しない点にある。ただし非言語情報には男女差や大人と子供の年齢差のようにほぼ全面的に音声器官の解剖学的特徴で決定されるものと、感情のように上述の声質にかかわるものとが認められるので、図14ではこれらも別の経路として認定している。

5 結論

音響学、生理学、心理学、言語学の観点からパラ言語情報の伝達メカニズムを検討し、その成果を概念モデルにまとめた。パラ言語情報の科学的研究は本研究その他によって、現在その端緒についたばかりである。数多くの重要な研究課題が今後の検討を待っている[.]。

参考文献

[1] H. Fujisaki, 1997. Prosody, Models, and Spontaneous Speech, In Y. Sagisaka, et al., (Eds.), *Computing Prosody: Computational Models for Processing Spontaneous Speech*, New York: Springer, 27-42..
 [2] 前川喜久雄, 「パラ言語的情報」別冊国文学「現代日本語必携」, No. 53, pp. 172-175 (2000).
 [3] K. Maekawa and T. Kagomiya, "Influence of paralinguistic information on segmental articulation," *Proc. 5th ICSLP2000*, Beijing, Vol. 2, pp.349-352 (2000).
 [4] 前川喜久雄、北川智利, 「パラ言語情報の知覚」日

本行動計量学会第28回大会論文集, pp.157-160 (2000).
 [5] 前川喜久雄、北川智利, 「音声はパラ言語情報をいかに伝えるか」*認知科学*, 9-1, pp.46-66 (2002).
 [6] M. Fujimoto and K. Maekawa, "Variation of phonation types due to paralinguistic information: An analysis of high-speed video images," *Proc. ICPhS 2004*, Vol. 4, pp.2401-24004 (2004).
 [7] O. Mizutani and N. Mizutani, *Aural Comprehension Practice in Japanese*. The Japan Times (1979).
 [8] 川上素, 「文末などの上昇調について」*国語研究*, No. 16, pp. 25-46 (1963).
 [9] 前川喜久雄, 「声質とパラ言語情報」, 韻律に着目した音声言語情報処理の高度化(科研費特定研究2003年度研究成果報告書), pp.15-18 (2004).
 [10] K. Maekawa, H. Kikuchi, Y. Igarashi and J. Venditti, "X-JToBI: An extended J_ToBI for spontaneous speech," *Proc. ICSLP 2002*, Denver, 3, pp. 1545-1548 (2002).
 [11] K. Maekawa. "Production and Perception of 'Paralinguistic' Information," *Proc. Speech Prosody 2004*, Nara, pp. 367-374 (2004).
 [12] 前川喜久雄, 「パラ言語情報研究の課題」*音響学会秋季研究発表会講演論文集*, pp.247-250 (2002).
 [13] Scherer, K. R., Cross-cultural investigation of emotion inferences from voice and speech: Implications for speech technology, *Proc. 5th ICSLP*, Beijing, 2, pp. 379-382 (2000).

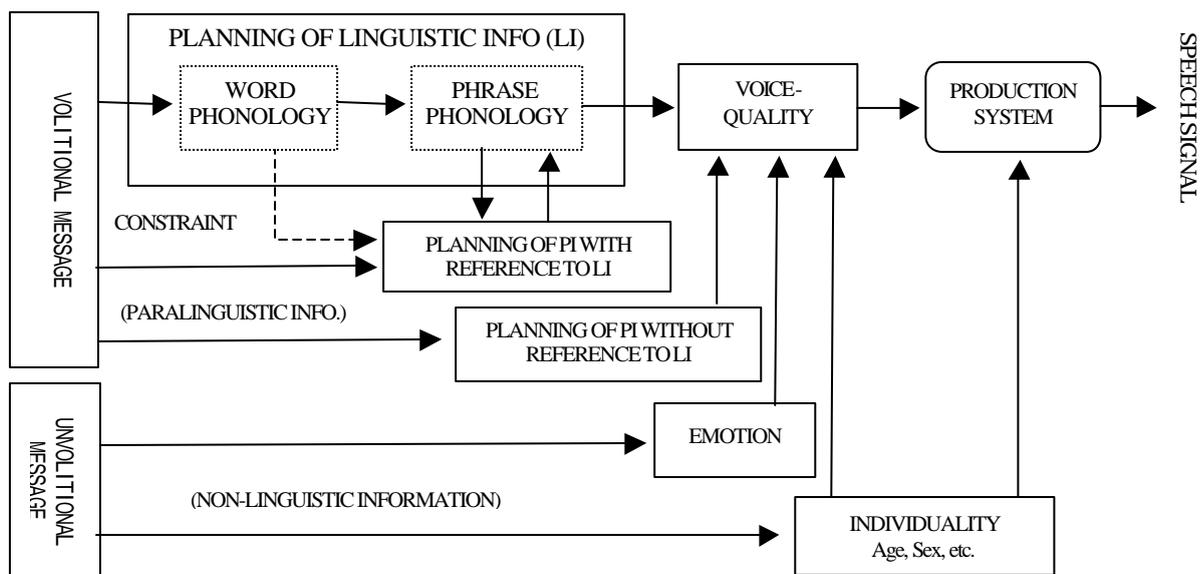


図14.パラ言語情報・非言語情報を考慮した音声生成過程の概念図